

2. Las funcionalidades del periodista en el actual panorama de divulgación de datos: carencias y nuevas habilidades

Journalist functionalities in the current data divulgation landscape: shortages and new skills

Alba Pereira Ferreira
alba.p.f.r@gmail.com
Universidad Complutense de Madrid

Resumen: Casi cualquier área de la vida diaria de los ciudadanos se ve impregnada de los usos e influencias de la datificación de sus hábitos, y por ello resulta cada vez más interesante poner a disposición del usuario la cantidad ingente de información generada, en un formato acorde al flujo informativo contemporáneo. Desde una perspectiva crítica, el propósito de este artículo es determinar la existencia o no, en el panorama vigente, de un Periodismo de datos real frente a un Periodismo con datos convencional, basado en la reproducción acrítica de un cuadro numérico de resultados que se toman como ciertos sin más averiguación y que, además, conllevan un trato deficiente de los mismos. Para ello, resulta necesario tomar en consideración la situación actual de acceso a los recursos y datos abiertos, así como la explotación de estos métodos. Tal y como afirman Mayer-Schonberger y Cukier (*Big Data. La revolución de los datos masivos*, 2013) “una vez que se ha datificado el mundo, los usos potenciales de la información no tienen más límite que el periodismo”. En este sentido, el periodismo señala cierta falta de compromiso con los datos, en especial en las regiones hispanoparlantes, y por ello se torna imprescindible analizar su reflejo en el ejercicio práctico y académico.

Al análisis de esta realidad acompaña el surgimiento de una nueva figura del profesional de la información en los medios de comunicación españoles, un híbrido de periodista que aúna el perfil de narrador de historias con el de diseñador y estadístico, aquél que analiza los datos y los expone de la forma más interesante. En suma, una máquina que, a partes iguales, filtra y reinterpreta los datos, incorporándoles el relato correspondiente y generando la visualización consecuente. La base de las anteriores teorías emplea una metodología de estudio triangular, basada en las entrevistas a expertos, la realización de cuestionarios a periodistas en activo dispersos en el horizonte mediático nacional y el análisis de contenido de medios que acreditan aportes significativos en la materia, tanto en el plano cuantitativo como cualitativo-interpretativo. En el transcurso del estudio se revelan posibles líneas de indagación futuras, relacionadas con la explotación noticiosa de recursos como el Small/Big Data, escaparate de grandes cantidades de información invisible a los ojos de la mayoría.

Palabras clave: Periodismo, datos, medios de comunicación, estadística, visualización.

Abstract: Almost any area of the daily life of citizens is impregnated with the uses and influences of their habit's datification. For that reason, it is becoming extremely interesting to make available for the users the huge amount of the generated information, in a format adapted to the contemporary informative flow. From a critical perspective, this article aims to determine, in the current landscape, the existence of a real data journalism compared to the conventional discipline, based on the uncritical reproduction of a numerical table of results, taken as certain without a further inquiry and that also involves a poor treatment of them. To prove that, it is necessary to consider the actual situation of the access to resources and open data, and the operation of these methods. As Mayer-Schonberger and Cukier claim (Big Data. The revolution of massive data, 2013) "once the world is datified, the potential uses of the information are not limited but only by the personal ingenuity." Although this reality is fulfilled, today's journalism notes a lack of commitment to data, especially in Spanish-speaking regions, and therefore it is essential to analyze its reflection in both practical and academic exercise.

In addition to the analysis of this reality, it is worth highlighting the emergence of a new professional figure in the Spanish media, an hybrid that combines the journalistic profile of an storyteller with the design and the statistic capabilities, who analyzes the data and exposes it in the most interesting way. In short, a machine that, in equal parts, filters and interprets the data, incorporating the appropriate story and generating the consequent display. The basis of the above theories, elementary on the analysis of data journalism, employs a triangular methodology study, based on interviews with experts, conducting questionnaires to active journalists dispersed in the national media and analysis of some media content accrediting significant contributions in the field, from a both quantitative and qualitative-interpretative perspective. Also during the study, possible future lines of inquiry, related to new resource exploitation as the Small / Big Data, arises.

Key words: Journalism, data, media, statistics, visualization

1. Introducción al estado de la cuestión. Evolución del Periodismo de datos vigente

El Periodismo de datos contemporáneo se vio especialmente marcado por dos acontecimientos que en 2012 marcaron una transición en la disciplina. Por un lado, el lanzamiento en enero de ese mismo año del primer concurso internacional centrado en la especialidad. La iniciativa, impulsada por Global Editors Network en alianza con Google supuso un antes y un después. Por otro lado, la publicación ese mismo año del primer manual de Periodismo de datos, ‘The data journalism handbook’ colaboró a su vez en este proceso (Mariño, 2012: 23).

Las iniciativas en torno al Periodismo de datos han comenzado siendo argumentadas en inglés, concretamente al otro lado del Atlántico, en Estados Unidos. El académico español José Luis Dader corrobora en sus estudios la teoría de que esta nueva versión del periodismo evolucionó de una forma mucho más lenta en Europa que en América del Norte y, dentro del Viejo Continente, todavía más despacio en España. Así lo declaraba en sus investigaciones como docente ya en el año 1993.

“En EE.UU., por el contrario, tal situación parece ya muy superada. La labor de preparación estadística en las escuelas de periodismo empieza a notarse y los principales medios cuentan con sus propios equipos especializados de realización y tratamiento de sondeos” (Dader, 1993: 104)

Aunque con el paso del tiempo esta disyuntiva se solucionó tan rápido como evolucionan hoy en día las nuevas Tecnologías de la Información y la Comunicación, para el autor España continuaba a la cola en los años 90. La razón principal era que los periodistas malinterpretaban y sacaban de contexto los resultados estadísticos que se deducían de los sondeos. A través del uso generalizado de las distintas herramientas conocidas en la actualidad, la constante reforma del Periodismo de datos y el Big Data (literalmente ‘grandes datos’ en español) se produjo de forma paralela en ambas regiones geográficas, propiciando, así, que en nuestros días existan iniciativas interesantes de este tipo de periodismo tanto en Europa en general como en España en particular. En concreto en el continente europeo, las primeras iniciativas de renombre en cuanto a Periodismo de datos se dieron en Inglaterra con *The Guardian* y en Alemania con *Der Spiegel*.

Para Dader, en esta implementación tardía en el mundo hispano tuvo mucho que ver el nivel de desarrollo en el ámbito de las ciencias en general. No en balde, el mundo anglosajón a lo largo de los últimos 150 años es el que ha avanzado a mayor velocidad en el mundo científico, generando una mentalidad por la que la gente valoraría más los datos y la importancia de la sistematización del conocimiento, influyendo en aspectos de la vida cotidiana de países como Estados Unidos, Gran Bretaña o Suecia, que consideran necesario tener buenas estadísticas de sus actividades. En cambio, en el mundo latino se ha pecado de ser más “relajado” y “espontáneo”. Todo ello formaría parte de un estrato cultural que ha influido en nuestro periodismo, más cercano a las humanidades, el arte, la literatura o la buena redacción, infravalorando elementos más propios de las ciencias duras.

Algunos hechos actuaron con anterioridad a los grandes hitos de 2012 como bisagra entre el avance y conocimiento del Periodismo de datos. Uno de ellos fue el llamado ‘Cablegate’, acontecido en el año 2010 y responsable de la filtración de un gran

número de documentos diplomáticos estadounidenses y de otras naciones, difundidos en el espacio en red Wikileaks. Tal y como revelaron los medios al poco tiempo con sus reacciones, que no se hicieron esperar, el Big Data había llegado para quedarse, atrayendo a la humanidad nuevos desafíos de poder y comunicación. Julian Assange, fundador de Wikileaks, escogió determinados diarios europeos para distribuir los documentos en el mundo occidental, siendo *El País* el elegido en España.

El periodista Paul Bradshaw vaticinó en su blog personal lo necesario de humanizar en la medida de lo posible el trato de grandes cantidades de datos: “hay que humanizar y personalizar grandes conjuntos de datos de una manera tal que ello no vaya en contra de la complejidad del tema que se está tratando” (Mariño, 2012: 23).

Lejos de ser una cuestión fácil, profundizar en esta masificación de los datos representa una gran complejidad, lo que hace necesarios a los periodistas para llevar a cabo ciertas funciones de recopilación. Todo ello justifica lo inevitable de replantear las cualidades profesionales y documentales del informador tal y como las hemos conocido hasta ahora.

Así lo refleja el análisis del periodista John Keenan, que en ese mismo año, 2010, afirmó en la cabecera *The Guardian* que “sin el análisis, el comentario y la explicación de periodistas especializados, los documentos habrían quedado en Wikileaks, llamando la atención sólo de aquellos con un ferviente interés en el conflicto. La Casa Blanca y el Pentágono, apenas se habrían inmutado”. Para los expertos, no hay duda de que se tiende a infravalorar la dependencia que se tiene del periodista, lo que representa es un mal de nuestro futuro más inmediato.

A la existencia de esta gran cantidad de información de la que venimos hablando acompaña una creciente necesidad de organización. Parece evidente que hoy en día las TIC producen datos más rápido de lo que crean conocimiento (Meyer, 2012). Y ahí es donde se fundamenta el papel del periodismo. Solo durante este año, más de un billón de gigabytes de datos nuevos se crearán a nivel mundial (Puyol, 2014: 472). El Big Data cuenta con un

gran espectro de posibles aplicaciones, que van desde las horas de video grabadas para vigilancia pública hasta los millones de ‘tweets’ que se emiten por día o segundo o los miles de seguimientos de los vehículos que se documentan por GPS. En la actualidad, la capacidad de almacenamiento de la Tierra alcanza los 2,5 zetabytes de información.

Disponiendo de este volumen en unidades de DVD, atesoraríamos una pila de discos tan grande que alcanzaría 1,5 veces la distancia hasta la luna (Puyol, 2014: 479). Según los expertos, esta cantidad crece aproximadamente a un ritmo del 50% cada año. En la antigüedad, cuando Tolomeo II de Egipto conservaba un ejemplar de cada obra escrita en la gran biblioteca de Alejandría, aquel almacén representaba la suma de todo el conocimiento del planeta (Mayer- Schönberger, 2013: 21). Con la revolución digital, el diluvio de datos existente equivaldría a darle a cada persona de la Tierra trescientas veinte veces la información (estimada) acumulada en aquella biblioteca.

Determinados estudiosos coinciden en que el proceso de digitalización al que se ve sometido el mundo global colabora en la desaparición de puestos de trabajo como el del periodista. Sin embargo, esta no es una visión generalizada de la cuestión. Muchos otros enfocan la problemática a favor del informador, de tal forma que, lejos de ser sustituido, el periodista se vuelve necesario en la síntesis y la interpretación del flujo constante de datos y, por consiguiente, de hechos (Meyer, 2012).

Resulta cada vez más palpable la evidencia, tal como reflexionaba el CEO de Vodafone en Reino Unido Guy Laurence en una entrevista con el periodista Simon Rogers, de que “antes el acceso a la información era un gran problema mientras que ahora, la prioridad es saber manejarla”. Philip Meyer reconocía en 2012 que eventualmente ya no sería suficiente con aplicar herramientas de estudio científico a los datos, sino que sería imprescindible una estructura para ver la verdad de los hechos.

“El periodista debe ahora ser un administrador de datos acumulados, un procesador de datos y un analista de esos datos. El proceso comprendería los siguientes pasos: recopilación, almacenaje, recuperación, análisis, resumen y comunicación o distribución”. (Meyer, 2012)

En épocas pasadas, el análisis de la información escasa de la que se disponía implicaba la necesidad del muestreo constante y un anhelo permanente de exactitud. Con la prevalencia de las tecnologías digitales de alto rendimiento, apreciamos detalles en grandes cantidades de datos, que nos ofrecen una vista despejada de lo concreto pudiendo asumir ciertos errores de medida y dejando de lado la rigidez en la investigación. Lo interesante es reconocer tendencias generales (Mayer- Schönberger, 2013: 26).

Aunque la exactitud deja de ser primordial, la calidad va en aumento con este nuevo modelo de la información. Meyer comparaba el desarrollo de la nueva era digital con el desarrollo de la agricultura. Cuando la caza y la agricultura eran las principales fuentes de alimento, las personas eran poco insistentes sobre la particularidad del producto que consumían. A medida que se vivió en la abundancia y se generó mayor cantidad de alimentos, naciendo la era industrial, los seres humanos comenzaron a preocuparse más por su contenido en vitaminas, su sabor, su textura, etc. Y lo mismo ocurre con la información.

En este sentido, para José Luis Dader lo importante es utilizar siempre las ciencias duras con rigor, evitando hacerles daño con incorrecciones o inexactitudes revestidas de cierta intencionalidad, ya que, algunas veces, ciertos fallos provocan distorsiones o que se publiquen noticias falsas con apariencia de exactitud. Así, finalmente, herramientas de estudio como los sondeos y las encuestas, que aventuran predicciones, se ven a menudo desacreditadas y descontextualizadas por ello.

Aunque el fenómeno de Wikileaks, tal y como mencionábamos antes, supuso una revolución en torno al Periodismo de datos, la disciplina ya había comenzado a despuntar de forma más pausada en los años ochenta y noventa, periodo en el que los sistemas informáticos se vuelven más accesibles para el gran público, otorgando a los periodistas la opción de recabar datos en lugares en los que antes no era posible. Hasta entonces, los métodos convencionales de recopilación de información eran las encuestas y los sondeos de opinión.

Con el paso de los años, el lanzamiento de nuevos medios, en especial en formato digital, ha motivado la necesidad de incrementar la labor de nuevos periodistas de datos que diesen valor a un tipo de ‘periodismo explicativo’ (Bounegru, 2014). Sitios de Estados Unidos y Reino Unido más allá de los archiconocidos proyectos de *The Guardian* y *The New York Times* encontraron su lugar en la nube velozmente. Algunos destacados fueron el blog *FiveThirtyEight*, del periodista Nate Silver, y la plataforma *Vox*, fundada por el blogger Ezra Klein. En concreto, el primer ejemplo ha llamado la atención notoriamente en los últimos tiempos con su fórmula de periodismo, dedicada en buena parte al relato como el complemento perfecto de sus números y enfocada hacia el aprendizaje del lector, que también conoce el proceso por el que se llega a cada historia. Para Josu Mezo, consultado en la elaboración de esta investigación, este modelo despertaría un gran interés, a pesar de ser difícilmente reproducible.

“Sus proyectos dan pistas para hablar con números sobre temas que inicialmente no tienen mucho que ver con ellos, resultando además bastante didácticos. Los periodistas se hacen preguntas que pueden parecer tontas o triviales, pero a medida que se explica cómo se llega a las historias y cómo se buscan se entiende la parte en que intervienen las estadísticas”.

Para el periodista James Ball, conocido por su estrecha relación con el Periodismo de datos en su colaboración con Wikileaks, existirían una serie de limitaciones comunes a los proyectos nacientes en torno a esta disciplina del oficio, yendo desde la falta de transparencia de sus informaciones hasta la perpetuación de la desigualdad de género en sus redacciones (“still a lot of white guys”) o la falta de relevancia actual de sus noticias (Bounegru, 2014).

El comunicador también señala que los periodistas recurren habitualmente a bases de datos y fuentes comunes de información que se encuentran preestablecidas, como las estadísticas de los gobiernos, evitando hacer el esfuerzo de encontrar o generar los datos por sí mismos. Es importante dar relevancia no solo al resultado final del análisis o a la erudición propia de los datos estadísticos, sino también a conocer de dónde procede la información, cómo es producida y cuáles son sus utilidades. Y ahí es donde se manifiesta el

principal problema y objeto de estudio: antes de averiguar cómo tratar con los grandes flujos de datos, es importante conocer hasta qué punto el dato, en sí mismo, va a recibir el tratamiento adecuado.

Mientras el periodista de datos promete un giro de la opinión hacia la evidencia, en ocasiones los datos no proceden tanto de su cosecha propia como de un pequeño número de fuentes establecidas. Por ello, Ball anima a los profesionales de este campo a buscar reflexiones frescas, nuevas, de una misma cuestión. Las fuentes disponibles van más allá de las instituciones y organizaciones conocidas, e incluso documentos sacados de la hemeroteca de nuestras bibliotecas pueden ayudarnos a elaborar una información nueva, desapercibida.

En la actualidad, un problema asociado al Big Data es su mitificación como fuente de información de forma automática. Para Michael Wu, científico de Analytics en Lithium, la gente a menudo no se da cuenta de que datos e información no siempre es lo mismo. Para el experto, es una realidad a menudo ignorada que "aunque seamos capaces de extraer información de grandes cantidades de datos, no toda será siempre lo bastante profunda y valiosa".

La diferencia entre ambos conceptos es bastante sutil. ‘Data’ no es más que un registro de eventos que tuvieron lugar, una descripción de algo que ha sucedido a través del cuándo, el dónde, el cómo y el quién. Algo informativo. Sin embargo, un mayor número de datos no siempre significa más información. Wu sostiene que cuantos más datos se tiene, menos información se gana proporcionalmente:

"The information you can extract from any big data asymptotically diminishes as your data volume increases. This does seem counterintuitive, but it is true". (Wu, 2012)

Para ilustrar lo anterior, resulta muy interesante el ejemplo que el científico da acerca de los ‘retweets’ de Twitter. Muchos ‘tweets’ son redundantes debido al ‘retweet’, y por ello el contenido de muchos de los mensajes que se emiten en la red social no son más

que copias y alusiones a cuestiones que ya se han tratado. También es una realidad que los mensajes lanzados por un mismo usuario tenderán a multiplicarse acumulando redundancias, ya que sus intereses y sus gustos se mantendrán en el tiempo, motivando la generación de contenido similar al existente con anterioridad.

Michael Wu determina: información \leq datos. Dentro de un océano de información disponible, gran parte de la muestra se revela difícilmente interpretable, y dentro de ese océano, un mar de datos interpretables se desecha por ser irrelevante o ya conocido, es decir, por ser lo que comúnmente se conoce como “ruido”. Así, llegamos a reconocer el valioso papel del narrador, que además guarda la inteligencia de saber extraer de ese mar lo que puede llegar a interesar de forma global.

2. El proceso de estudio. Análisis de contenido, cuestionario y entrevistas a expertos

En el proceso investigativo que da origen al presente artículo, realizamos un estudio metodológico de base triangular, apoyado en los pilares que describimos a continuación: el análisis de contenido de varios medios digitales y en torno a un espacio temporal acotado; la realización de entrevistas a expertos, cuyas opiniones sirvieron como apoyo a las hipótesis iniciales y las conclusiones, y, por último, el análisis de cuestionarios sobre el ejercicio del Periodismo de datos en las redacciones españolas remitidos a periodistas en activo.

En suma, la investigación se ha centrado en el análisis tanto en el plano cuantitativo como en el cualitativo-interpretativo de muestras escogidas dentro de un periodo previamente acotado. Con ello, se ha buscado la identificación y selección de todas las informaciones publicadas en las que se desplegasen un número significativo de datos cuantificados y comentados. Se incluyeron informaciones relativas a ámbitos diversos, incluyendo géneros periodísticos dispares como el reportaje o la noticia.

Desarrollar la investigación en la Comunidad de Madrid facilitó en gran medida el acceso a personalidades experimentadas en el plano de la datificación, masiva o no masiva. Así, las principales fuentes de contrastación fueron el académico y experto en periodismo

de precisión José Luis Dader, un auténtico referente en la materia en España; David Cabo, informático y propulsor de la plataforma *Civio* de datos abiertos; Laura Tejedor Fuentes, doctora en transparencia democrática, y Josu Mezo, sociólogo y estadístico dedicado en sus ratos libres al análisis del tratamiento erróneo de los datos en los medios de comunicación españoles.

Se realizó el análisis de contenido generado por cinco medios digitales españoles escogidos, por un lado, por su relevancia en el panorama informativo actual y, por otro, por la amplia disponibilidad de sus archivos o hemerotecas, de las que otros títulos carecen, y que facilitan mucho el poder recuperar días completos para su posterior estudio. Son *Elpaís.es*, *Elmundo.es*, *eldiario.es*, *El Blog de El Español* y *El Confidencial*.

El periodo escogido fue de 30 días naturales, repartidos entre los meses de junio y julio de 2015, de enorme interés en España a nivel político y económico por, entre otras razones, la activación de la precampaña electoral de cara a las elecciones generales de diciembre y el reflejo de los nuevos datos del paro y las variaciones del empleo en la temporada de verano.

Aunque inicialmente se había considerado el análisis de las cabeceras *El País* y *El Mundo* en su versión en papel, tras realizarse un mínimo seguimiento de estos títulos durante algunos días se comprobó que la actividad en ellos, en lo que a datos se refería, era prácticamente nula, por lo que se optó por desecharlos de la metodología. Cobra mucho más sentido analizar diarios y blogs online que, por un lado, aglutinan la mayoría de lectores habituales, y, por otro, explotan las posibilidades reales de la visualización interactiva.

A su vez, se llevó a cabo el análisis de los últimos post del blog de Josu Mezo, *Malaprensa*, y el espacio radiofónico *El Purgatorio* entre los meses de julio y agosto para extraer algunas conclusiones en torno al periodismo “con” datos, que no “de” datos, que venimos tratando, subiendo un nivel para analizar los datos que no solo se exponen sin

filtros o investigación, sino que además se extraen erróneamente de la matriz que los contiene, generando un periodismo “con” datos maltratados.

Entre los errores más comunes encontramos, por ejemplo, el de sumar varios porcentajes calculados desde diferentes bases para hallar una media. Los periodistas suelen apoyar sus errores con las cifras en que ‘son de letras’, afirmación a la que Mezo responde que los fallos que cometen poco tienen que ver con las matemáticas o estadística.

En relación a los preguntados en el cuestionario aplicado y a la tipología y la denominación del medio al que pertenece cada uno, nos encontramos con que participan tanto prensa escrita y digital como medios audiovisuales, agencias y gabinetes de comunicación, sumando en total 20 profesionales que responden a nuestras cuestiones entre el 14 de abril y el 20 de julio de 2015. Así, destacan en prensa *eldiario.es*, *ideal.es*, *infobierzo.es*, *Diario de Pontevedra*, *El Norte de Castilla* y *La Voz de Galicia*. De radio y televisión colaboran *Ahora Granada* y *Radio Televisión Canal Sur Radio*. En cuanto agencias, participan periodistas de la *Agencia EFE*, y, de gabinetes, periodistas de la Secretaría de Estado de Comunicación del Ministerio de la Presidencia y del gabinete de comunicación de la ONG *WWF*.

3. Conclusiones. El maltrato de los datos y la concepción errónea del periodista multidisciplinar

En el desarrollo de las conclusiones, cabe realizar una división en consideración de cada una de las partes analizadas. En primer lugar, en torno a la configuración de las funcionalidades actuales del periodista en las redacciones. En segundo lugar, en base al perfil observado del periodismo de precisión y en torno a sus posibilidades en el campo del análisis de los datos.

Resulta acertado llegar a la conclusión de que en la actualidad se llama Periodismo de datos no es más que el antiguo Periodismo de precisión, bautizado por Philip Meyer en sus inicios hace ya casi medio siglo. Aunque la terminología ha cambiado, se trata del

mismo tipo de periodismo, una variante de la disciplina dedicada al estudio de bases de datos en la historia contemporánea.

Laura Tejedor nos recuerda que, tal y como señala Mar Cabra, y contra lo que dicta la opinión pública, desconocedora de los antecedentes de esta modalidad de reportajes o noticias, el Periodismo de datos está de moda pero no es una moda, y lo reflejan numerosos acontecimientos del ámbito periodístico acaecidos en las últimas décadas.

En relación a los cuestionarios realizados, y en torno a la importancia del acceso a los datos abiertos, merece la pena destacar que para la muestra de periodistas encuestada, la percepción es que, a pesar de que Open Data implica una apertura de la información disponible, desde la Administración se plantean restricciones que limitan la eficiencia del periodista. En general, la filosofía Open Data ha supuesto una revolución de alcance global, por la que cada vez accederemos a una mayor cantidad de información. Sin embargo, a nivel de país, los recursos de datos abiertos han surgido siempre con cierta dificultad y rodeados por la polémica, dificultando en muchas ocasiones la labor del informador.

En relación a las funciones del periodista, y a la posible actualización de su perfil en un entorno de análisis y recopilación de datos, cabe concluir considerando estos cuestionarios, pero también la opinión de nuestros expertos entrevistados.

Teniendo en cuenta la muestra contactada, no representativa del ámbito periodístico pero sí de interés para la presente investigación, se puede decir que el perfil profesional mayoritario en las redacciones es el del redactor de nivel bajo o medio, encargado de la recopilación y la publicación de la información. En relación a los estudios y el nivel de formación de los encuestados, resulta que en el 87,5% de los casos se trata de periodistas o comunicadores audiovisuales, que suelen encontrarse inmersos en la realización de cursos de formación complementarios o estudiando herramientas digitales de forma autodidacta.

Es oportuno deducir que, extrapolando el resultado de los formularios, aproximadamente la mitad de los periodistas en activo han recibido directrices o apoyo en torno al uso de aplicaciones de las TIC en su trabajo, si bien mayoritariamente esta formación responde a una necesidad del propio medio, habitualmente desarrollado en el entorno digital. No es una medida baja, teniendo en cuenta que determinados periodistas, en especial los de más edad y con una trayectoria dilatada en la profesión, carecen de estos conocimientos. En este sentido, podríamos establecer una clara diferencia entre una y otra generación de periodistas, teniendo en cuenta que los más jóvenes han desarrollado sus estudios universitarios en el ámbito de las nuevas tecnologías.

En cuanto al uso de recursos multimedia, como el texto, la imagen, el vídeo o el sonido, la creación propia o de autoría mixta (compartida entre el propio medio y recursos ajenos) alcanza el 32,9% de la muestra. En especial, el uso de la visualización, de gran interés para alcanzar determinadas conclusiones en el presente estudio, se revela de autoría propia o relativa al medio en el 52,9% de los casos, lo que nos permite llegar a la conclusión de que la cantidad de medios interesados en este ámbito podría ser en la esfera nacional de aproximadamente la mitad.

Haciendo un compendio de las respuestas registradas en relación al proceso de la investigación en el periodismo en general, llegamos a la determinación de que el 75% de la muestra encuestada sabe dar respuestas coordinadas cuando se le pregunta por el significado de conceptos de reciente aparición como ‘Big Data’ u ‘Open Data’, o de otros con más historia, como ‘Periodismo de precisión’, de modo que, en general, las nuevas generaciones de periodistas se encuentran predispuestas a profundizar en su conocimiento.

Así, se puede aventurar que, aunque la mayoría de periodistas no desarrolla funcionalidades específicas relacionadas con los datos en su ámbito de trabajo, conoce sus implicaciones y utilidades teóricamente, predisponiéndose a su investigación y desarrollo. Algo más de la mitad de la muestra (el 64,7%) se considera capacitada para la lectura de estadísticas y su elaboración, valorando, además, muy útil el uso de la infografía para representar resultados. Frente a esta percepción positiva de sus capacidades, los

periodistas se exponen a menudo al ojo crítico de expertos en la materia, como Josu Mezo, que consideran que lo que parece algo evidente en estadística para algunos, a menudo, no lo es tanto.

Teniendo en cuenta todo lo anterior, y redundando en la idea de que los periodistas se encuentran capacitados desde su conocimiento teórico para profundizar en materia de datos, cuando llegamos a la práctica la cuestión vira de forma inesperada. En base al análisis de la última cuestión planteada en este sentido, percibimos que los periodistas preguntados carecen de la lógica o la capacidad necesaria para ordenar los elementos ‘datos/filtrado/visualización/historia’¹, lo que revela un problema de base.

A pesar de que la muestra es pequeña y las conclusiones derivadas han de ser tomadas con cautela, es natural pensar que, si nuestro objetivo es construir una historia, debemos partir de los datos desestructurados, para filtrarlos consecuentemente y aplicar una visualización que nos ayude a elaborar el relato posterior.

Resulta revelador que ninguno de los encuestados ordenase correctamente esta enumeración, considerando, además, que para muchos la base es una historia ya existente, la filtración si queremos llamarlo así. Para expertos como David Cabo, es bien cierto que, en el entorno de datos masivos en el que nos movemos, impone o “da miedo” pensar en el desconocimiento que demuestran los periodistas.

Para Dader, ambas fórmulas de periodismo (la que parte de la investigación y la que parte de datos que ya han sido trabajados) son aceptables, siempre y cuando se aplique el análisis necesario. Desgraciadamente, y tal y como venimos observando en el análisis de contenido, este no suele ser el caso.

En cuanto al perfil del infografista o del diseñador gráfico en las redacciones, solamente el 31,2% de los medios encuestados cuenta con esta clase de expertos en su

¹ Evolución del ‘Data-driven journalism’ según el periodista Mirko Lorenz.

medio de trabajo, de lo que se deriva que, en el resto, o son los periodistas los que asumen estas funciones, o no se trabaja apenas en los datos, ya que a este tipo de análisis suele acompañar la visualización. El porcentaje se reduce hasta el 12,5% cuando hablamos de expertos estadísticos.

Teniendo en cuenta todo lo anterior, revisando la opinión de los expertos que han participado en el estudio y analizando el contenido de los medios digitales elegidos, recuperamos la idea inicial de la que parte el estudio, centrada en las nuevas funciones del periodista en el entorno de los datos, para refutar su argumento, ya que la realidad nos indica que no asistimos al nacimiento de una nueva figura del periodista, uniendo su perfil narrativo al del estadístico, el investigador, el analista de datos o el diseñador gráfico.

Muy pocos profesionales de la comunicación son capaces en la actualidad de desarrollar con éxito el conjunto de estas vertientes, que, sin embargo, son necesarias si queremos llevar a cabo un reportaje de datos de interés. Por ello, se deduce que, ante la imposibilidad de crear un híbrido de periodista multiusos, las redacciones han de hacerse cada vez más multidisciplinares, tratando de reunir en un mismo espacio de trabajo todos estos perfiles.

Según corrobora José Luis Dader, “el concepto del ‘periodista total’ está bien en cuanto a la versatilidad del profesional, pero está mal en cuanto a su capacidad para tomar un tema a fondo para tratarlo con rigor y dedicarle el tiempo necesario”. Por ello, aunque un periodista puede desarrollar nociones básicas de ciertas disciplinas para enriquecer su trabajo y su visión del mismo, lo cierto es que si diversifica sus funciones difícilmente podrá dedicarle a cada una la profundización que requiere.

Es bien cierto que, aunque estas áreas pueden ser muy accesibles, son casos contados los de quienes las dominan. Un ejemplo es el Josu Mezo que, dedicado a la estadística y a su enseñanza desde una perspectiva sociológica, se interesa por el periodismo y el uso que éste hace de aquello que él conoce bien. David Cabo comparte este

punto de vista, considerando que aunque “él es algo excepcional, algo raro. Te desmonta muchas noticias, pero lo hace como estadístico interesado en el periodismo”.

Investigar y ahondar en los datos implica muy a menudo saber programar y, para el ingeniero, la moda de que el periodista aprenda a hacerlo significa confiar en un modelo utópico. Se puede concluir que lo es pensar que un periodista narrativo va a destinar la otra mitad de su cerebro a este tipo de cuestiones técnicas. Así, lo más factible es considerar el Periodismo de datos como un trabajo de equipo. “Un periodista es muy, muy difícil que aprenda a programar siendo bueno también escribiendo. Sería un unicornio. Es muy útil, sin embargo, que tenga nociones, ya que no tiene por qué ser un total ignorante en estadística y en programación”. Así sabrá lo que es factible, lo que es interesante, y podrá compararlo con otras cosas ya hechas.

Laura Tejedor coincide en esta percepción, señalando que incluso en los grandes consorcios internacionales de Periodismo de datos se dan las coaliciones entre periodistas e informáticos. “Es el caso de ‘España en llamas’, de *Civio* (David Cabo). En él colaboraron periodistas como Marcos García Rey, pero también ingenieros como Juan Elosua. Es un buen ejemplo de esa simbiosis que reclamamos desde el Periodismo de datos, porque siempre se necesita un informático”.

Incluso Josu Mezo, desde su punto de vista técnico, considera que, aunque es plausible aprender casi de todo, lo frecuente no es que lo haga un periodista. Avanzando en nuestro análisis, podemos confirmar el potencial que implica el análisis de datos masivos o Big Data, ampliamente demostrado y contrastado por nuestras fuentes, así como la necesidad de los recursos abiertos, tanto de software como de informaciones, para alcanzar el Periodismo de datos real. Sin embargo, no se cumple la expectativa en España, y esta es la parte que debemos reseñar.

Llegamos así a concluir en este apartado haciendo uso de la terminología que hemos empleado a lo largo del estudio para distinguir el periodismo “de” datos, que encuentra su principio y su final, su razón de ser, en los datos, y el periodismo “con” datos, que,

llanamente, se basa en el relato apoyado en cifras “masticadas” por terceros, como complemento.

Aunque la información disponible para la generación de contenido noticiable se multiplica por dos cada año, algunos expertos, como David Cabo, consideran que todavía nos encontramos en una fase “burbujil” del Big Data, valorando que su potencial es infinito pero no explotado en su plenitud, más todavía en el ámbito del periodismo.

Tras la columna de humo que ha surgido en los últimos años como resultado de su descubrimiento, la búsqueda y análisis de cantidades masivas de datos sigue estando algo estancada. Es, además, una realidad que los periodistas no disponen del tiempo ni la economía necesaria para dedicarse a la elaboración de grandes reportajes con los datos. Por esta y otras razones, surgen dos problemáticas: la de llamar Periodismo de datos a algo que no es más que periodismo con cifras o números, facilitados por entidades e instituciones dedicadas al estudio social y económico de la realidad, y la del maltrato de esos datos, incluso cuando llegan ya trabajados.

La primera problemática planteada ha sido analizada en la primera pata de nuestra metodología con el análisis de contenido de las versiones digitales de *El país.es*, *El mundo.es*, *eldiario.es*, *El Confidencial* y *El Blog de El Español*.

La realidad es que únicamente en *El Confidencial* se hace una labor cercana al tratamiento de los datos, aunque la mayoría de las veces se construye un relato en base a fuentes oficiales o institutos de estadística que, con cierta periodicidad, elaboran estudios noticiables. También es *El Confidencial* el único medio que trabaja activamente en la visualización, explotando herramientas actuales como ‘Tableau’ en la elaboración de gráficos interactivos interesantes, uno de los principales atractivos de esta clase de periodismo.

Lo anterior viene a confirmar los testimonios de nuestros entrevistados expertos, que corroboran, sin excepción, el hecho de que este periódico digital es el único que en

España “tiene un equipo dedicado a revisar estas cosas”. Según Laura Tejedor, este medio fue, además, pionero en crear en España una unidad de datos inserta en su redacción. “*El Confidencial* es el único que trabaja en proyectos no tan sujetos a la actualidad, con un margen de preparación muy delicado, que permite sacar adelante trabajos muy atractivos visualmente”.

En el caso del resto de medios consultados, *El Blog de El Español* destaca por publicar con cierta periodicidad trabajos con los datos de considerable interés, lo que puede avanzar hasta convertirse en una de sus señas identificativas.

El resto no revelaron dedicación en el asunto, ni en pequeña ni en gran medida, ya que las únicas alusiones a los datos se hacen como parte de informaciones diarias sujetas a la rigurosa actualidad y apoyadas en estudios oficiales que responden a intereses particulares, generando, por tanto, periodismo “con” datos. Es curioso el caso de *Eldiario.es*, que en el periodo escogido prácticamente no trató los datos, aún a pesar de que supuestamente cuenta con un laboratorio integrado por varios miembros. Uno de ellos participó en el cuestionario del presente estudio.

Tanto *El País* como *El Mundo* demuestran tener diseñadores gráficos integrados en su plantilla, y por ello a menudo trabajan en sencillas visualizaciones que, sin embargo, no constituyen Periodismo de datos en sí mismas.

En cuanto al maltrato de los datos, dentro del tratamiento de las cifras ya trabajadas con anterioridad por terceros, se produce una lectura muy a menudo errónea de las mismas, con una exposición nada acertada, tanto en la visualización como en el relato, tal y como afirma el sociólogo Josu Mezo. Así se corrobora en el análisis de contenido de *Malaprensa* y *El Purgatorio*.

Entre los errores más comunes, el de la falta de sentido común y el de la ausencia de la aplicación de herramientas básicas para la contrastación de las informaciones, de lo que se deduce la violación de uno de los principios fundamentales del periodismo, con o sin

datos, “no te fíes únicamente de tu primera fuente”. Como resultado, se proporcionan datos hallados mediante métodos inapropiados.

Para Mezo, el periodismo actual adolece en su base de la copia acrítica de los datos, ya que “muchas veces los periodistas tienen miedo a escarbar o cuestionar los números, afectados por el estigma de que supuestamente son de letras”. A esto acompaña la no activación de conocimientos de nivel usuario en materias como las matemáticas: “Cuando doy charlas a periodistas siempre insisto mucho en que los errores de los que hablo son errores con números, pero no de matemáticas o de estadística propiamente. A veces tienen que ver con la no comprensión de lo que significan los números, más que con que no sepan operar con ellos”.

Por lo tanto, es real que los datos no son adecuadamente utilizados ni explotados en el periodismo español. Que esta disciplina es relevante nos lo demuestra la datificación masiva a la que asistimos en los tiempos actuales, al igual que su empleabilidad en otras áreas como el comercio o el marketing, en las que se revelan de enorme utilidad para conocer a un público potencial.

Únicamente algunos consorcios de periodistas internacionales dedicados a los datos cuentan con la huella del investigador español, que si tiente trabajar en ello dentro de su país debe responder a organizaciones sin ánimo de lucro como *Civio*, cuyo principio y fin es estructurar los datos para una población que por sí misma no los comprende.

De todo lo anterior, se desprende que los medios de comunicación españoles se encuentran años luz de explotar adecuadamente las posibilidades noticiosas de recursos como el análisis de datos masivos, dejando abierta la posibilidad de retomar su estudio en profundidad en los próximos años, tal y como respalda David Cabo: “(En torno a Big Data aplicado al periodismo) por ahora prácticamente no se ha hecho nada, en especial en España. Se trata de algo potencial, futurible en un par de años, pero de momento teórico”.

4. Bibliografía y relación de fuentes

L Bounegru (2014). “What Data Journalist Need to Do Differently” en *Harvard Business Review*. Disponible en <https://hbr.org/2014/05/what-data-journalists-need-to-do-differently/> [Consulta: 30 de mayo de 2015]

JL Dader et P Gómez (1993). “‘Periodismo de precisión’ Una nueva metodología para transformar el periodismo” en *Anàlisi. Quaderns de Comunicació i Cultura*, pp. 99/116. Disponible en <http://ddd.uab.cat/pub/analisi/02112175n15/02112175n15p99.pdf> [Consulta: 28 de diciembre de 2014]

M Lorenz (2010) *Data-driven journalism: What is there to learn?* en *European Journalism Centre* (EJC). Disponible en http://mediapusher.eu/datadrivenjournalism/pdf/ddj_paper_final.pdf [Consulta: 19 de marzo de 2015]

V Mayer- Schönberger et K Cukier (2013). *Big Data. La Revolución de los datos masivos*. Madrid: Turner Publicaciones S.L. Disponible parcialmente en <https://books.google.es/books?hl=es&lr=&id=uO9FbEcaMpkC&oi=fnd&pg=PA11&dq=big+data:+la+revolucion+de+los+datos+masivos&ots=VYAT1ilKAX&sig=AKa9pVKiOJDsMrDjfkWej1PJFGM#v=onepage&q=big%20data%3A%20la%20revolucion%20de%20os%20datos%20masivos&f=false> [Consulta: 28 de diciembre de 2014]

A Mariño (2012). “La relación del Periodismo de datos con Big Data y Open Data”. Proyecto Final de Máster. Madrid: Universidad Carlos III. Disponible en <https://es.scribd.com/doc/111830154/La-relacion-del-Periodismo-de-Datos-con-Big-Data-y-Open-Data#scribd> [Consulta: 15 de diciembre de 2014]

P Meyer (2012). “Precision Journalism and Narrative Journalism: Toward a Unified Field Theory en *Nieman Reports*. Disponible en <http://niemanreports.org/articles/precision-journalism-and-narrative-journalism-toward-a-unified-field-theory/> [Consulta: 3 de abril de 2015]

J Puyol (2014). "Una aproximación a Big Data" en *Revista de Derecho UNED*, núm. 14. Disponible en <<http://revistas.uned.es/index.php/RDUNED/article/view/13303/12174>> [Consulta: 3 de abril de 2015]

M Wu (2012). "The Big Data Fallacy and Why We Need To Collect Even Bigger Data" en *TechCrunch*. Disponible en <<http://techcrunch.com/2012/11/25/the-big-data-fallacy-data-%E2%89%A0-information-%E2%89%A0-insights/>> [Consulta: 9 de julio de 2015]

Alba Pereira Ferreira es Licenciada en Periodismo por la Universidad de Valladolid (2014) y actualmente doctoranda del Programa de Periodismo de la Universidad Complutense de Madrid. Ha desarrollado distintas actividades como comunicadora, pasando entre otros por el Diario de Pontevedra como redactora, y también desde la comunicación institucional donde trabajó en el Ministerio de la Presidencia.